# Analysis of the Convergence Properties of Self-Normalized Source Separation Neural Networks

Yannick Deville, *Member, IEEE*

***Abstract*— An extended source separation neural network was recently derived by Cichocki *et al.* from the classical Hérault–Jutten network. It claimed to have several advantages, but its convergence properties were not described. In this paper, we first consider the standard version of this network. We determine all its equilibrium points and analyze their stability for a small adaptation gain. We prove that the stationary independent sources that this network can separate are the globally sub-Gaussian signals. As the Hérault–Jutten network applies to the same sources, we thus show that the advantages of the new network are not counterbalanced by a reduced field of application, which confirms its attractiveness in the considered conditions. Moreover, we then introduce and analyze a modified version of this network, which can separate the globally super-Gaussian source signals. These theoretical results are experimentally confirmed by computer simulations. As a result of our overall investigation, a method for processing each one of the two classes of signals (i.e. sub- and super-Gaussian) is available.**

## I. Introduction

Blind source separation is a generic signal (and data) processing technique that applies, e.g., to antenna or microphone array processing [1]. In its "simplest configuration," two measured signals $x_1(t)$ and $x_2(t)$ are available (e.g., from sensors), and these signals are unknown linear instantaneous mixtures of two unknown independent source signals $s_1(t)$ and $s_2(t)$, i.e. (see Fig. 1)

$$x_1(t) = a_{11}s_1(t) + a_{12}s_2(t) \quad (1)$$

$$x_2(t) = a_{21}s_1(t) + a_{22}s_2(t) \quad (2)$$

where the terms $a_{ij}$ are unknown real-valued constant (nonzero[1]) mixture coefficients such that[2] $a_{11}a_{22} - a_{12}a_{21} \neq 0$. The source separation problem then consists of estimating the source signals $s_j(t)$ from the measured signals $x_i(t)$ up to a permutation and scaling factor. Although the linear instantaneous mixture model is simple, the separation of a set of sources from such mixed signals has various applications. Several of them concern radio-frequency systems (e.g., satellite communications [2], protection against garbling in radar applications [3], multitag identification systems [4], [5]), but others are, e.g., related to the analysis of signals measured in mechanical structures (movements of dams [6], nondestructive control of the generators of nuclear power plants [6]).

The first two solutions to the linear instantaneous source separation problem were proposed by Bar-Ness *et al.* [2] and Hérault and Jutten (see, e.g., [1], [7], and [8]) at the beginning of the 1980's. Since then, source separation has been a very active research field, and many alternative approaches have been developed. For a survey of all these methods, refer to [9]. For the sake of brevity, we only consider hereafter the methods directly related to the specific approach analyzed in this paper, i.e., we focus on the solutions initially proposed by Hérault and Jutten and then extended by other authors.

The Hérault–Jutten approach is based on a recurrent artificial neural network. The convergence properties of this network have been analytically studied by various independent authors in the "simplest configuration" defined above. Sorouchyari [10] and Fort [11] used almost the same method, which was then revisited and somewhat extended by Moreau and Macchi [12], [13]. Comon *et al.* [14] presented another method that yields different results. An approach bridging the gap between these two methods was then proposed by Deville [15] so that the convergence properties of the simplest versions of this network are now well defined.

Two classes of structures related to the Hérault–Jutten network were then also proposed for performing linear instantaneous source separation. On the one hand, Moreau and Macchi [12], [13], [16] introduced a direct (i.e., nonrecurrent) version of the Hérault–Jutten network based on the same adaptation rule. This network is attractive because it avoids the matrix inversion that must be performed with the recurrent version in order to derive the outputs from the inputs and network weights. Moreau and Macchi also studied the convergence properties of this network in the "simplest configuration," using the same type of method as Sorouchyari and Fort. They also proposed and studied a mixed version of this network [12], [13].

On the other hand, Cichocki *et al.* [17] defined neural networks that may be considered to be extensions of the

Manuscript received December 3, 1996; revised August 6, 1998. The associate editor coordinating the review of this paper and approving it for publication was Dr. Sergios Theodoridis.

The author is with the Laboratoire d'Acoustique, Métrologie, Instrumentation (LAMI), Toulouse, France (e-mail: ydeville@cict.fr).

Publisher Item Identifier S 1053-587X(99)03234-1.

[1] The case when at least one mixture coefficient $a_{ij}$ is zero is not detailed here for two reasons. On the one hand, it is of no importance in this paper because it implies that at least one of the measured signals $x_1(t)$ and $x_2(t)$ is not a mixed signal so that the initial source separation configuration reduces to a classical adaptive filtering (or gain control) problem, which should be treated with a simpler signal processing structure than the one considered here. On the other hand, not considering that useless case somewhat simplifies the stability analysis, as shown in Section IV.

[2] The latter condition corresponds to the invertibility of the mixing matrix $A$ consisting of the mixture coefficients $a_{ij}$. This condition is required; otherwise, the two measured signals are identical (up to a scalar factor). In the latter case, the mixing matrix cannot be inverted by using these measured signals, meaning that the source separation problem cannot be solved (no separating equilibrium points with finite weight values exist in this case, as shown in the demonstration of Theorem 2).

above-mentioned ones. These new networks contain additional self-adaptive weights, which are updated to normalize the "scales" of the network outputs. These networks were claimed to be able to process ill-conditioned mixtures and badly scaled source signals to which the Hérault-Jutten network would not apply. Both the direct and recurrent versions of these types of neural networks were described, and it was also proposed to cascade them in a multilayer neural network in order to improve performance. To our knowledge, the papers published up to now only describe the principles of these networks as well as their empirical performance derived from numerical simulations. On the contrary, no theoretical proof has been provided about their convergence properties. This may result from the fact that such analyses are significantly more complex than for the Hérault–Jutten and Moreau–Macchi networks, as will be shown in this paper, due to the normalization scheme introduced by Cichocki in his networks.

Therefore, at the current stage, the Cichocki networks seem to be more powerful than the simpler formerly proposed structures, but despite these advantages, we are entitled to hesitate to use them, as their behavior is not well defined. It is especially not known whether these networks have spurious stable equilibrium points, that is, weight values toward which they may converge without providing separated signals at their outputs. Similarly, it is not known whether they are able to separate a limited or large class of source signals. Such restrictions have been shown to exist for the Hérault–Jutten and Moreau–Macchi networks. Similar limitations are therefore expected to arise with the Cichocki networks, and they should be determined. This paper therefore aims at precisely defining the conditions of operation of these networks. It should first be noted that the overall properties of the multilayer versions of these networks are directly derived from those of the individual layers that compose them because each of these layers is adapted by a local adaptation rule, i.e., independently from the other layers. Therefore, the convergence analysis only needs to be performed for the single-layer versions of these networks. Moreover, only the direct version of these networks is considered hereafter since it is more attractive than the recurrent one as explained above and because similar results are expected for both versions based on the similarity observed between the properties of the Hérault–Jutten and Moreau–Macchi networks.

The remainder of this paper is organized as follows. The considered type of source separation networks is defined in Section II. Two specific versions of these networks are then studied. As they yield similar analyses, only the investigation of the standard version is reported in detail in the subsequent sections, whereas the modified version that we introduce for processing other types of source signals is described more rapidly in Appendix D. More precisely, the presentation of the standard version contains the following aspects. Its equilibrium points are studied in Section III. The stability of these points is then investigated in Section IV. The theoretical results thus obtained are checked by means of computer simulations in Section V. Eventually, Section VI presents the conclusions drawn from this investigation and outlines its potential extensions.

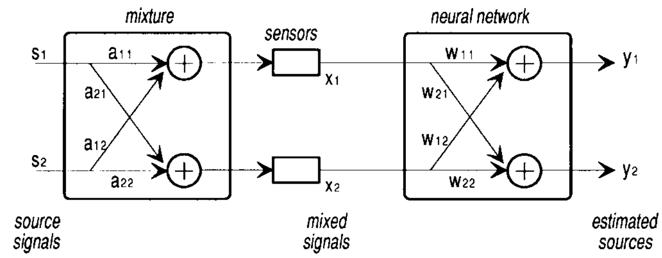


Fig. 1. Basic source separation configuration and direct Cichocki network.

## II. Definition of the Considered Class of Networks

This section briefly describes the principles of the class of direct single-layer Cichocki networks which is analyzed in the subsequent sections. This investigation is performed for the "simplest configuration" defined in Section I, and the source signals $s_1(t)$ and $s_2(t)$ are assumed to be stationary, zero-mean, and statistically independent. At each time step $t$, the network (shown in Fig. 1) receives both mixed signals $x_1(t)$ and $x_2(t)$ defined by (1) and (2) and processes them as follows.

1) It computes its output signals $y_1(t)$ and $y_2(t)$ corresponding to the current input signals and internal weight values $w_{ij}$, i.e.,[3]

$$y_1(t) = w_{11}x_1(t) + w_{12}x_2(t) \tag{3}$$
$$y_2(t) = w_{21}x_1(t) + w_{22}x_2(t). \tag{4}$$

2) It also updates its four real-valued weights $w_{ij}$ according to the following adaptation rules:

$$\frac{dw_{ii}(t)}{dt} = -\mu[f[y_i(t)]g[y_i(t)] - 1], \quad i \in \{1,2\} \tag{5}$$
$$\frac{dw_{ij}(t)}{dt} = -\mu f[y_i(t)]g[y_j(t)], \quad i \neq j \in \{1,2\} \tag{6}$$

where $\mu$ is a small positive adaptation gain, and $f$ and $g$ are distinct odd functions called the "separating functions" below.

The principles of the adaptation rules (5) and (6) may be summarized as follows:

- Rule (6) is used for updating the cross-coupling weights $w_{12}$ and $w_{21}$. It aims at making the outputs independent and thus, respectively, proportional to each source. It is the same rule as that of the Hérault–Jutten and Moreau–Macchi networks.
- Rule (5) is used to update the direct weights $w_{11}$ and $w_{22}$. It aims at normalizing the "scales" of the output signals $y_1(t)$ and $y_2(t)$, respectively. The scaling coefficients $w_{11}$ and $w_{22}$ are meant to self-adapt to the processed signals. This rule is specific to the Cichocki network (whereas in the Hérault–Jutten and Moreau–Macchi networks, $w_{11}$ and $w_{22}$ are fixed to 1). The above-mentioned advantages of the Cichocki network result from the adaptation of these weights.

[3] The weight values $w_{ij}$ depend on $t$. For readability, this is often omitted in the notations used below.

The last parameters that must be defined in order to fully specify the considered version of the network are the selected separating functions $f$ and $g$. These parameters appear in the Cichocki, Hérault–Jutten, and Moreau–Macchi networks. Various functions have been considered in the papers related to all these networks. The most commonly used set of functions is [1], [4], [5], [10]–[13], [15]–[17]

$$f(x) = x^3 \quad \text{and} \quad g(x) = x. \tag{7}$$

The subsequent sections only concern this specific set, whereas another set is considered in Appendix D.

It should be noted that the "neural" nature of the Hérault–Jutten, Moreau–Macchi, and Cichocki networks does not lie in their input/output transfer function. Whereas a specific feature shared by many neural networks consists in transferring input signals through nonlinear (most often sigmoidal) functions, the networks considered in this paper have a linear transfer function. Instead, their neural nature resides in their original inspiration [8] and in the nonlinearity used in their adaptation rules, i.e., in their nonlinear separating functions $f$ and/or $g$.

## III. Equilibrium Points

### A. Definition of the Equilibrium Points

The equilibrium points of the network adaptation rules (5) and (6) are all the quadruplets of constant weight values $(w_{11}, w_{12}, w_{21}, w_{22})$ such that

$$\left\langle \frac{dw_{ij}(t)}{dt} \right\rangle = 0, \quad i, j \in \{1, 2\} \tag{8}$$

where $\langle\,\rangle$ stands for mathematical expectation. Condition (8) may be rewritten by using (5) and (6), showing that the equilibrium points are the solutions of

$$\langle f[y_i(t)]g[y_i(t)]\rangle = 1, \quad i \in \{1, 2\} \tag{9}$$

$$\langle f[y_i(t)]g[y_j(t)]\rangle = 0, \quad i \neq j \in \{1, 2\}. \tag{10}$$

Especially for the separating functions defined in (7), these equations become

$$\langle [y_i(t)]^4\rangle = 1, \quad i \in \{1, 2\} \tag{11}$$

$$\langle [y_i(t)]^3 y_j(t)\rangle = 0, \quad i \neq j \in \{1, 2\}. \tag{12}$$

Given the source signals $s_j(t)$ and mixture coefficients $a_{ij}$, the problem is then to derive the corresponding quadruplets of weight values that meet (11) and (12). To this end, we first determine the corresponding composite matrices $P$ defined as follows (see also [17]). The mixture equations (1) and (2) may be rewritten in matrix form as

$$X(t) = AS(t) \tag{13}$$

where $X(t) = [x_1(t), x_2(t)]^T$, $S(t) = [s_1(t), s_2(t)]^T$, and $A$ is the mixture matrix composed of the mixture coefficients $a_{ij}$. Similarly, (3) and (4), defining the transfer function of the network, may be rewritten in matrix form as

$$Y(t) = WX(t) \tag{14}$$

where $Y(t) = [y_1(t), y_2(t)]^T$, and $W$ is the weight matrix composed of the weight values $w_{ij}$. From (13) and (14), the overall relationship between the source signals $s_j(t)$ and the output signals $y_i(t)$ may be written in matrix form as

$$Y(t) = PS(t) \tag{15}$$

or explicitly as

$$y_1(t) = p_{11}s_1(t) + p_{12}s_2(t) \tag{16}$$

$$y_2(t) = p_{21}s_1(t) + p_{22}s_2(t) \tag{17}$$

where $P$, which is composed of the real-valued elements $p_{ij}$, is defined as

$$P = WA. \tag{18}$$

By using (16) and (17) and taking into account the independence of the source signals, the equilibrium conditions (11) and (12) may be rewritten as[4]

$$p_{11}^4\langle s_1^4\rangle + 6p_{11}^2p_{12}^2\langle s_1^2\rangle\langle s_2^2\rangle + p_{12}^4\langle s_2^4\rangle = 1 \tag{19}$$

$$p_{21}^4\langle s_1^4\rangle + 6p_{21}^2p_{22}^2\langle s_1^2\rangle\langle s_2^2\rangle + p_{22}^4\langle s_2^4\rangle = 1 \tag{20}$$

and

$$p_{11}^3p_{21}\langle s_1^4\rangle + 3p_{11}p_{12}(p_{12}p_{21} + p_{11}p_{22})\langle s_1^2\rangle\langle s_2^2\rangle + p_{12}^3p_{22}\langle s_2^4\rangle = 0 \tag{21}$$

$$p_{21}^3p_{11}\langle s_1^4\rangle + 3p_{21}p_{22}(p_{22}p_{11} + p_{21}p_{12})\langle s_1^2\rangle\langle s_2^2\rangle + p_{22}^3p_{12}\langle s_2^4\rangle = 0. \tag{22}$$

The resolution of these four equations provides the expressions of the four elements $p_{ij}$ of the composite matrix for each equilibrium point with respect to the moments of the source signals. The determination of these elements $p_{ij}$ is presented in the following subsections. The corresponding weight values may then be derived from (18) (see the beginning of Appendix B).

### B. Equilibrium Points with at Least One Zero Element $p_{ij}$

As a first step, we determine the four specific sets of solutions of (19)–(22) that, respectively, correspond to each of the additional conditions $p_{11} = 0$, $p_{12} = 0$, $p_{21} = 0$, or $p_{22} = 0$. Straighforward computations show that the condition $p_{11} = 0$ implies $p_{22} = 0$ and vice versa so that each of the two conditions $p_{11} = 0$ and $p_{22} = 0$ leads to the same solution, corresponding to both $p_{11} = 0$ and $p_{22} = 0$. Similarly, the conditions $p_{12} = 0$ and $p_{21} = 0$ yield the same solution. Eventually, the solutions of (19)–(22) for which at least one element $p_{ij}$ is zero are the following ones:

- a first set containing four solutions (depending on the selected values of $\epsilon_1$ and $\epsilon_2$)

$$p_{11} = \frac{\epsilon_1}{\langle s_1^4\rangle^{1/4}}, \quad \epsilon_1 \in \{-1, 1\} \tag{23}$$

$$p_{12} = 0 \tag{24}$$

$$p_{21} = 0 \tag{25}$$

$$p_{22} = \frac{\epsilon_2}{\langle s_2^4\rangle^{1/4}}, \quad \epsilon_2 \in \{-1, 1\}. \tag{26}$$

[4] In the remainder of this paper, the time index $t$ is omitted in the mathematical expectations of signals.

These equilibrium points are denoted $F_{\epsilon_1,\epsilon_2}$ hereafter. For each of them, (16) and (17) become

$$y_1(t) = p_{11}s_1(t) \tag{27}$$
$$y_2(t) = p_{22}s_2(t). \tag{28}$$

Therefore, at these points, each network output $y_i(t)$ is proportional to $(\propto)$ the source having the same index $i$, i.e., the network achieves the separation of the source signals without permuting them. These solutions only differ in the signs of the coefficients by which the source signals are multiplied.

- a second set that also contains four solutions

$$p_{11} = 0 \tag{29}$$
$$p_{12} = \frac{\epsilon_2}{\langle s_2^4\rangle^{1/4}}, \quad \epsilon_2 \in \{-1, 1\} \tag{30}$$
$$p_{21} = \frac{\epsilon_1}{\langle s_1^4\rangle^{1/4}}, \quad \epsilon_1 \in \{-1, 1\} \tag{31}$$
$$p_{22} = 0. \tag{32}$$

These equilibrium points are denoted $G_{\epsilon_1,\epsilon_2}$ hereafter. For each of them, (16) and (17) become

$$y_1(t) = p_{12}s_2(t) \tag{33}$$
$$y_2(t) = p_{21}s_1(t). \tag{34}$$

Therefore, at these points, $y_1(t) \propto s_2(t)$ and $y_2(t) \propto s_1(t)$, i.e., the network achieves the separation of the source signals while permuting them. Again, these solutions only differ in the signs of the coefficients by which the source signals are multiplied.

### C. Equilibrium Points with Only Nonzero Elements $p_{ij}$

The only solutions of (19)–(22) that were not provided by the method used in the previous subsection are those that meet the additional requirement

$$p_{ij} \neq 0, \quad i,j \in \{1, 2\}. \tag{35}$$

Their determination is detailed in Appendix A. It yields the following eight points[5] (depending on the selected values of $\epsilon_1$, $\epsilon_2$ and $\epsilon_3$):

$$p_{11} = \frac{\epsilon_3}{\left[2\langle s_1^4\rangle\left(1 + 3\frac{\langle s_1^2\rangle\langle s_2^2\rangle}{\sqrt{\langle s_1^4\rangle\langle s_2^4\rangle}}\right)\right]^{1/4}}, \quad \epsilon_3 \in \{-1, 1\} \tag{36}$$
$$p_{12} = \frac{\epsilon_2 p_{11}}{R}, \quad \epsilon_2 \in \{-1, 1\} \tag{37}$$
$$p_{21} = \epsilon_1\epsilon_2 p_{11}, \quad \epsilon_1 \in \{-1, 1\} \tag{38}$$
$$p_{22} = \frac{-\epsilon_1 p_{11}}{R} \tag{39}$$

with

$$R = \left(\frac{\langle s_2^4\rangle}{\langle s_1^4\rangle}\right)^{1/4}. \tag{40}$$

[5] Additional solutions exist in the specific case of globally Gaussian sources, as explained in Appendix A.

These equilibrium points are denoted $H_{\epsilon_1,\epsilon_2,\epsilon_3}$ hereafter. For all of them, the outputs of the network are still expressed according to (16) and (17) with (35). Therefore, at these points, the network outputs mixtures of the two sources, thus failing to achieve source separation.

### D. Discussion

To be able to solve the source separation problem defined in Section I, the system under investigation should have equilibrium points at which it outputs separated sources. The above description shows that the network considered in this paper indeed meets this requirement (the corresponding points are the eight points $F_{\epsilon_1,\epsilon_2}$ and $G_{\epsilon_1,\epsilon_2}$, which are called the "separating equilibrium points" below). However, it also has "spurious" equilibrium points, i.e., equilibrium points at which it outputs mixtures of the sources (namely, the eight points $H_{\epsilon_1,\epsilon_2,\epsilon_3}$, which are called the "nonseparating equilibrium points" below). We must then determine which of these 16 equilibrium points are stable, depending on the nature of the sources, or at least which types of sources can be separated by this network as a result of some stability properties of these equilibrium points. That is the topic of the next section.

## IV. Stability of the Equilibrium Points

### A. Tangent Mean Algorithm and First Stability Conditions

*1) General Principles:* Sorouchyari [10] and Fort [11] used the same approach to investigate the local stability of the equilibrium points of the Hérault–Jutten network. This method was then also applied by Moreau and Macchi [12], [13], [16] (with a reference to the "ordinary differential equation" technique [18]) to their modified versions of the Hérault–Jutten network.[6] For any fixed separating functions[7] and equilibrium point $E$, this method consists of i) considering the mean adaptation algorithm, here obtained by replacing the right side of (5) and (6) by its mathematical expectation and ii) deriving a first-order development of this mean algorithm at point $E$. This yields, in vector form

$$\frac{d\Delta V}{dt} = J\Delta V \tag{41}$$

where $J$ is the Jacobian matrix of the system at point $E$. $\Delta V$ is the column vector defined as $\Delta V = [\Delta w_{11}, \Delta w_{12}, \Delta w_{21}, \Delta w_{22}]^T$, where each component $\Delta w_{ij}$ is the difference between two values of $w_{ij}$, respectively, at a considered point in the neighborhood of $E$ and at point $E$ itself. A necessary and sufficient condition for point $E$ to be locally stable is then as follows: C1) the real parts of all the eigenvalues of $J$ are negative [10]. An equivalent condition is C2), which is the real parts of all the roots of the characteristic polynomial $Q(\Lambda)$ associated with $J$ are negative, where $Q(\Lambda)$ is defined as

$$Q(\Lambda) = \det(J - \Lambda I). \tag{42}$$

[6] Moreau and Macchi also considered related approaches targetted at the discrete-time and stochastic versions of the adaptation rule [12], [13], [16].

[7] Although the overall method may be applied to any separating functions, it has been detailed only for specific functions in the literature, i.e., especially for (7).

It should be clear that this approach is accurate only for a vanishing adaptation gain $\mu$ because the first-order development that it uses yields a linearized approximation of the considered network around the selected equilibrium point. On the contrary, for a nonvanishing adaptation gain, this approach only provides approximate results.

The first step of the method used in this paper is based on the principles defined above and consists of computing the Jacobian matrix $J$ and the associated characteristic polynomial $Q(\Lambda)$ of the considered network, as shown in the remainder of this subsection. However, the next step, consisting of the practical exploitation of condition C2), leads to a specific approach, which is presented in Sections IV-B2 to IV-B5.

It should be noted that the adaptation rule of the considered network has some similarities with the class of constant modulus adaptive (CMA) algorithms. Therefore, the analysis of its mean behavior presented below has some links with the investigations concerning CMA algorithms reported in the literature (see, especially, the capture analysis presented in [19]).

*2) Application to the Considered Network:* When using the separating functions (7), relatively long but straightforward calculations yield

$$J = rM \tag{43}$$

with

$$r = -\frac{\mu}{D} \tag{44}$$

$$D = w_{11}w_{22} - w_{12}w_{21} \tag{45}$$

and

$$M = \begin{bmatrix} 4w_{22} & -4w_{21} & 0 & 0 \\ -3w_{12}\langle y_1^2 y_2^2\rangle & 3w_{11}\langle y_1^2 y_2^2\rangle & w_{22} & -w_{21} \\ -w_{12} & w_{11} & 3w_{22}\langle y_1^2 y_2^2\rangle & -3w_{21}\langle y_1^2 y_2^2\rangle \\ 0 & 0 & -4w_{12} & 4w_{11} \end{bmatrix} \tag{46}$$

where $w_{ij}$ are the weight values at the considered equilibrium point $E$. Hence, (42) allows the derivation that

$$Q(\Lambda) = r^4 P(\lambda) \tag{47}$$

with

$$\Lambda = r\lambda \tag{48}$$

$$P(\lambda) = p_4\lambda^4 + p_3\lambda^3 + p_2\lambda^2 + p_1\lambda + p_0 \tag{49}$$

$$p_4 = 1 \tag{50}$$

$$p_3 = -(w_{11} + w_{22})\big(3\langle y_1^2 y_2^2\rangle + 4\big) \tag{51}$$

$$p_2 = \langle y_1^2 y_2^2\rangle^2 9w_{11}w_{22} + \langle y_1^2 y_2^2\rangle 12\big(w_{11}^2 + w_{22}^2 + 2D\big) + 15w_{11}w_{22} \tag{52}$$

$$p_1 = -4(w_{11} + w_{22})D\big[9\langle y_1^2 y_2^2\rangle^2 + 12\langle y_1^2 y_2^2\rangle - 1\big] \tag{53}$$

$$p_0 = 16D^2\big[9\langle y_1^2 y_2^2\rangle^2 - 1\big]. \tag{54}$$

Equations (47) and (48) allow us to link the roots $\Lambda_i$ of $Q(\Lambda)$ to the roots $\lambda_i$ of $P(\lambda)$, according to

$$\Lambda_i = r\lambda_i. \tag{55}$$

The polynomial $P(\lambda)$ is therefore used as an intermediate variable in the following analysis.

*B. Discussion of Stability Conditions*

Based on C2), the next natural step of the method would consist of

i) determining the analytical expressions of the roots of $P(\lambda)$;
ii) deriving the expressions of the roots of $Q(\Lambda)$ [by using (55)] as well as of their real parts;
iii) investigating the signs of these real parts, depending on the considered equilibrium point $E$ and on $\langle y_1^2 y_2^2\rangle$ [and, hence, on the source statistics due to (16) and (17)].

This approach could be used for the Hérault–Jutten network because its associated polynomial $P(\lambda)$ is only of order 2 so that it yields simple computations.[8] From a theoretical point of view, this method also applies to the network considered in the current paper because $P(\lambda)$ is of order 4, and the analytical expressions of its roots can be determined using Ferrari's method [20]. This approach is, however, very impractical as the expressions of the roots that we derived for the polynomial corresponding to (49)–(54) turn out to be very complicated.

An alternative approach can be developed by using the following principles. In fact, from the stability point of view, there is no need to determine the exact expressions of the roots of $Q(\Lambda)$. Instead, we only need a necessary and sufficient condition under which the real parts of all these roots are negative, i.e., under which $Q(\Lambda)$ is a Hurwitz polynomial [20]–[22]. This type of problem has been studied in the literature [20]–[22] and yields the following necessary and sufficient stability condition for any fourth-order polynomial $Q(\Lambda)$[9]

$$\left.\begin{aligned} & q_3 > 0 \\ & \begin{vmatrix} q_3 & q_4 \\ q_1 & q_2 \end{vmatrix} > 0 \\ & \begin{vmatrix} q_3 & q_4 & 0 \\ q_1 & q_2 & q_3 \\ 0 & q_0 & q_1 \end{vmatrix} > 0 \\ & q_0 > 0 \end{aligned}\right\} \tag{56}$$

where the parameters $q_i$ are the coefficients of $Q(\Lambda)$, here derived from the coefficients $p_i$ of $P(\lambda)$ by using (47) and (48). This approach is attractive because it directly provides a complete stability condition (56) for each equilibrium point, unlike the method outlined at the beginning of this subsection. However, developing this condition explicitly for the polynomial $Q(\Lambda)$ corresponding to (47)–(54) yields some complicated expressions. From these expressions, we cannot easily carry out the last step of the stability analysis, i.e., gather the stability conditions at all equilibrium points to

[8] The Hérault–Jutten network gives rise to a second-order polynomial $P(\lambda)$ because it only contains two adaptive weights (since $w_{11}$ and $w_{22}$ are fixed to 1).

[9] This description applies to the case when $q_4 > 0$. This condition is met here because $q_4 = p_4 = 1$.

eventually derive which types of sources can be separated by the considered network.

Therefore, in addition to the theoretically complete but impractical solution based on (56) that we already defined above, we developed an original approach that we describe hereafter. This method consists of focusing on some specific stability properties that are sufficient for determining which types of sources can be separated by the network operating with the separating functions (7).

### C. Necessary Stability Condition

*1) General Principles:* The proposed approach uses the fact that a simple necessary condition for any given equilibrium point to be stable can be derived very easily. This condition is based on the following theorem, which holds for any fourth-order polynomial $P(\lambda)$ with $p_4 = 1$.

*Theorem 1:* For any equilibrium point, a necessary stability condition is C3) $p_0 > 0$.

*Proof:* If $P(0) < 0$, $P(\lambda)$ has at least one real positive and one real negative root because $P(\lambda)$ is a continuous function of $\lambda$ and $P(\lambda) \to +\infty$ when $\lambda \to \pm\infty$ as $p_4 = 1$, and so has $Q(\Lambda)$, whatever the sign of $r$, due to the relationship (55) between the roots of $P(\lambda)$ and $Q(\Lambda)$. C2) is, therefore, not met. Similarly, if $P(0) = 0$, one of the roots of $P(\lambda)$ and, therefore, of $Q(\Lambda)$ is zero. C2) is, therefore, not met. In other words, stability requires $P(0) > 0$. Moreover, $P(0) = p_0$. This yields Theorem 1. □

Note that this result can also be derived from (56). Each of the four conditions contained by (56) is a necessary stability condition, especially the last one, i.e., $q_0 > 0$, which is equivalent to $p_0 > 0$ since (47) and (48) yield $q_0 = r^4 p_0$.

*2) Application to the Considered Network:* When using the separating functions (7), $P(\lambda)$ is defined by (50)–(54). Theorem 1 then shows that for any equilibrium point, a necessary stability condition is

$$\langle y_1^2 y_2^2 \rangle > \frac{1}{3}. \tag{57}$$

This condition should preferably be expressed with respect to the source signals $s_j(t)$. To this end, the moment $\langle y_1^2 y_2^2 \rangle$ should first be expressed with respect to these source signals. This depends on the relationship between the source signals and output signals $y_i(t)$ and, therefore, on the considered equilibrium point.

- For any of the equilibrium points $F_{\epsilon_1,\epsilon_2}$, (23)–(28) yield

$$\langle y_1^2 y_2^2 \rangle = \frac{\langle s_1^2 \rangle \langle s_2^2 \rangle}{\sqrt{\langle s_1^4 \rangle \langle s_2^4 \rangle}}. \tag{58}$$

  Condition (57) then reads

$$\frac{\langle s_1^2 \rangle \langle s_2^2 \rangle}{\sqrt{\langle s_1^4 \rangle \langle s_2^4 \rangle}} > \frac{1}{3}. \tag{59}$$

- All the equilibrium points $G_{\epsilon_1,\epsilon_2}$ also yield (58) and (59).
- For any of the nonseparating equilibrium points $H_{\epsilon_1,\epsilon_2,\epsilon_3}$, (16), (17), and (36)–(39) yield

$$\langle y_1^2 y_2^2 \rangle = \frac{1 - \frac{\langle s_1^2 \rangle \langle s_2^2 \rangle}{\sqrt{\langle s_1^4 \rangle \langle s_2^4 \rangle}}}{1 + 3\frac{\langle s_1^2 \rangle \langle s_2^2 \rangle}{\sqrt{\langle s_1^4 \rangle \langle s_2^4 \rangle}}}. \tag{60}$$

  Condition (57) then reads

$$\frac{\langle s_1^2 \rangle \langle s_2^2 \rangle}{\sqrt{\langle s_1^4 \rangle \langle s_2^4 \rangle}} < \frac{1}{3}. \tag{61}$$

### D. Interpretation of the Necessary Stability Condition

When using the separating functions (7), we showed above that for any equilibrium point, the necessary stability condition (57) is equivalent to (59) or (61). Therefore, it only depends on the considered type of sources (and, e.g., not on the values of the mixture coefficients $a_{ij}$). This stability condition can then be reinterpreted by splitting the analysis according to the possible types of sources, instead of using the partition related to the types of equilibrium points, which was considered in Section IV-C2. This yields the following results.

- The first case corresponds to the union of globally super-Gaussian and Gaussian sources, where the globally super-Gaussian sources are defined [11] as the sources that meet (61), and the globally Gaussian sources are those such that [11]

$$\frac{\langle s_1^2 \rangle \langle s_2^2 \rangle}{\sqrt{\langle s_1^4 \rangle \langle s_2^4 \rangle}} = \frac{1}{3}. \tag{62}$$

  In this case, (59) is not met; therefore, all the separating equilibrium points are unstable. Therefore, the network cannot converge to any of these points so it fails to achieve source separation. This completes our stability analysis for such sources. The remainder of our analysis, therefore, only concerns the other type of sources.
- The second case corresponds to the globally sub-Gaussian sources, which are defined [11] as the sources that meet (59) and are encountered, e.g., in [4] and [5]. In this case, (61) is not met; therefore, all the nonseparating equilibrium points are unstable. Therefore, in this case, the network cannot converge to any of these undesired points. With regard to the separating equilibrium points, no conclusions can be drawn about their stability yet, as (59), although it has been met here, is only a necessary condition. Therefore, two situations are *a priori* possible:
  - —Situation 1: All the separating equilibrium points are unstable. In that case, the network cannot converge to any of these points and, therefore, fails to separate the sources.
  - —Situation 2: At least one of the separating equilibrium points is stable. The latter then being the only stable points, the network can only converge to such a point[10] and thus succeeds in separating the sources.

[10] This may require an adequate initialization point and a low enough adaptation gain $\mu$ for the network to remain in the attraction domain of this point.

In order to determine whether the situation that actually occurs is "situation 1" or "situation 2," additional properties are used. They are described in the subsequent subsection.

### E. Definition of a Stable Separating Equilibrium Point

As explained above, we restrict ourselves here to the case when the separating functions (7) are used and the sources are globally subGaussian. The last step of the stability analysis is then based on the following theorems, the proofs of which are, respectively, provided in Appendices B and C.

*Theorem 2:* For the separating functions (7), if all mixture coefficients $a_{ij}$ are nonzero, at least one of the separating equilibrium points is such that $w_{11} > 0$, $w_{22} > 0$, and $D > 0$, where $D$ is defined in (45).

*Theorem 3:* For the separating functions (7) and for globally sub-Gaussian sources, if a separating equilibrium point is such that $w_{11} > 0$, $w_{22} > 0$, and $D > 0$, it is stable.

Theorem 2 shows the existence of a specific type of separating equilibrium points (when all $a_{ij}$ are nonzero), whereas Theorem 3 shows its stability in the considered conditions. As an overall result, the network has at least one stable separating equilibrium point, i.e., the situation that actually occurs is "situation 2" defined above. As explained in Section IV-D, the network is therefore able to separate all the considered sources, i.e., all the globally sub-Gaussian sources and none of the other types of sources as was proved in Section IV-D, which completes the stability analysis.

## V. Simulation Results

In order to illustrate the theoretical results obtained in the previous sections, we performed various simulations with the considered version of the network. The simulation conditions are defined hereafter, and their results are then described.

### A. Simulation Conditions

*1) Source Signals:* Each source $S_i(t)$ used in the simulations was a random binary-value signal, taking the values $+1$ and $-1$, respectively, with the probabilities $p$ and $(1-p)$. In each simulation, the same value $p$ was used for both sources. This value was varied over the simulations. It should be noted that these sources are not zero-mean, except for $p = 1/2$, as

$$\langle S_i(t) \rangle = 2p - 1. \tag{63}$$

The network was modified accordingly, using the approach already reported [8] for the Hérault–Jutten network. Briefly, (3) and (4) were applied to the actual nonzero-mean signals $Y_i(t)$ and $X_j(t)$, and (5) and (6) were used with estimates of the zero-mean versions $y_i(t)$ of the outputs $Y_i(t)$. The stability conditions thus apply to the zero-mean sources $s_i(t)$ corresponding to the actual sources $S_i(t)$, i.e.,

$$s_i(t) = S_i(t) - \langle S_i(t) \rangle. \tag{64}$$

It can be easily shown that for the sources $s_i(t)$ considered here, the global sub-Gaussianity condition (59) is equivalent to

$$p_l < p < p_h \tag{65}$$

where the lower and higher bounds $p_l$ and $p_h$ are equal to

$$p_l = \frac{1}{2} - \frac{\sqrt{3}}{6} \simeq 0.21 \tag{66}$$

$$p_h = \frac{1}{2} + \frac{\sqrt{3}}{6} \simeq 0.79. \tag{67}$$

Similarly, for such sources, the global super-Gaussianity condition (61) is equivalent to

$$p < p_l \quad \text{or} \quad p > p_h. \tag{68}$$

To ease the interpretation of the simulation results presented in the next subsection, a comment should be made here about the graphical representations of the considered sources. The representation of one such source consists of a discontinuous set of points having two $Y$-coordinate values, i.e., the two possible values of the source. These points are visible when the source is represented over a short time period. However, they appear as two horizontal lines when a long time period is considered.

*2) Mixture Parameters* All simulations were performed with artificial mixed signals $X_i(t)$ created as linear instantaneous combinations of the above-defined source signals $S_j(t)$. In other words, the mixed signals were derived by using the mixture equations (1) and (2), where all zero-mean signals were replaced by the actual nonzero-mean signals $X_i(t)$ and $S_j(t)$. The following mixture coefficients were used in all simulations:

- $a_{11} = 1$;
- $a_{12} = 0.8$;
- $a_{21} = 0.4$;
- $a_{22} = 1$.

Again, it should be noted that each mixed signal has four possible values so that it appears as four horizontal lines when represented over a long time period.

### B. Simulation Results

Simulations were performed for various values of the probability $p$ defining the sources. The following results were thus obtained.

*1) Simulations with Sub-Gaussian Sources:* When $p$ is chosen so that (65) is met, the network succeeds in separating the sources. This is illustrated in Figs. 2 and 3 for the case $p = 0.5$. Fig. 2 represents the evolution of the output signal $Y_1(t)$, which contains the following phases (the output signal $Y_2(t)$ yields the same results).

- The network weights start from 0, and so does $Y_1(t)$, due to (3).
- The first 4000 samples correspond to the network convergence phase. The four curves that appear in Fig. 2 during this phase result from the fact that $y_1(t)$ then contains a mixture of the two sources. This corresponds to the phenomenon described above for mixed signals, except that the curves are slanted here instead of being straight and horizontal, as the magnitude of the weights and, therefore, of $y_1(t)$ increases during this phase.
- After about 4000 samples, the weights have converged to constant values so that the magnitude of $y_1(t)$ remains

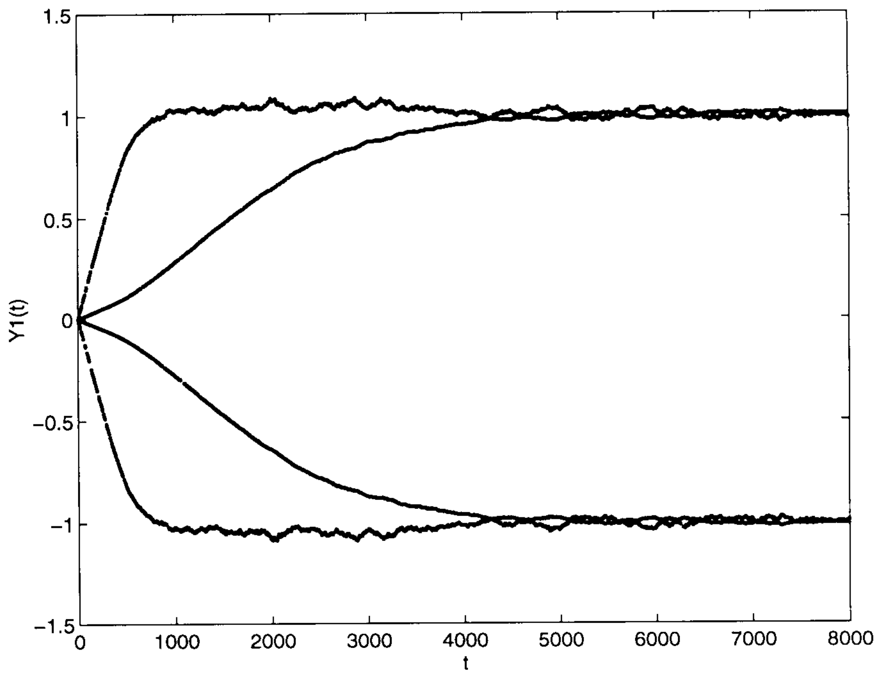


Fig. 2. Network output signal $Y_1(t)$ for sources corresponding to $p = 0.5$ (discrete set of points: see text).

constant. Moreover, this signal then appears as only two lines, which means that it only contains a single source (like in the above description of source signals, as opposed to that of mixed signals).

Fig. 3 represents the evolution of the network weights. This illustrates the considerations about these weights provided above and makes it possible to analyze the results more quantitatively as follows. The general expressions of the weight values at all separating equilibrium points are provided in Appendix B. Hereafter, they are applied to the specific sources and mixture coefficients considered here, taking into account the expressions of the terms $p_{ij}$ provided in Section III. Denoting

$$C = a_{11}a_{22} - a_{12}a_{21} \tag{69}$$

the weights at any of the four equilibrium points $F_{\epsilon_1,\epsilon_2}$ are expressed as

$$w_{11} = \frac{a_{22}}{C} p_{11} \simeq 1.47\epsilon_1 \tag{70}$$

$$w_{12} = \frac{-a_{12}}{C} p_{11} \simeq -1.18\epsilon_1 \tag{71}$$

$$w_{21} = \frac{-a_{21}}{C} p_{22} \simeq -0.59\epsilon_2 \tag{72}$$

$$w_{22} = \frac{a_{11}}{C} p_{22} \simeq 1.47\epsilon_2 \tag{73}$$

with $\epsilon_1,\ \epsilon_2 \in \{-1, 1\}$. Similarly, the weights at any of the four equilibrium points $G_{\epsilon_1,\epsilon_2}$ are expressed as

$$w_{11} = \frac{-a_{21}}{C} p_{21} \simeq -0.59\epsilon_2 \tag{74}$$

$$w_{12} = \frac{a_{11}}{C} p_{21} \simeq 1.47\epsilon_2 \tag{75}$$

$$w_{21} = \frac{a_{22}}{C} p_{21} \simeq 1.47\epsilon_1 \tag{76}$$

$$w_{22} = \frac{-a_{12}}{C} p_{21} \simeq -1.18\epsilon_1. \tag{77}$$

Comparing these values with Fig. 3 confirms that the network does converge to one of the separating equilibrium points and, more precisely, to $F_{1,1}$. It should be noted that this provides an experimental validation of Theorem 3. At point $F_{1,1}$, (70) shows that $w_{11} > 0$, (73) shows that $w_{22} > 0$, and $D = (Cw_{11}w_{22})/(a_{11}a_{22}) > 0$. As $p = 0.5$ results in (65), i.e., in sub-Gaussian sources, Theorem 3 allows us to conclude that $F_{1,1}$ is stable, i.e., that the network can converge to this point; this actually occurs, as shown above.

*2) Simulations with Super-Gaussian Sources:* When $p$ is chosen so that (68) is met, the network converges, but to a nonseparating equilibrium point. This is illustrated in Fig. 4 for the case $p = 0.1$. This figure yields the same comments as Fig. 2, except for the following aspects. When the network has converged to a nonseparating equilibrium point, the output signal $Y_1(t)$ is still a mixed signal. Therefore, it has four possible values, corresponding to four horizontal lines, as explained above. Moreover, the situation considered here is a degenerated case, where two of these four values become equal. This may be explained as follows. Here, the same value $p$ is used for both sources so that $\langle s_1^4 \rangle = \langle s_2^4 \rangle$. The expressions of the terms $p_{ij}$ provided in Section III then show that at any given nonseparating equilibrium point $H_{\epsilon_1,\epsilon_2,\epsilon_3}$, corresponding to a fixed value $\epsilon_1$

$$y_1(t) = p_{11}[s_1(t) + \epsilon_1 s_2(t)]. \tag{78}$$

As $s_1(t) = \pm 1$ and $s_2(t) = \pm 1$, the possible values of $y_1(t)$ are then $2p_{11}$, 0, and $-2p_{11}$. In other words, two of the four values that are different in the general case here both become equal to 0. The convergence to a nonseparating point then corresponds to a signal represented as three horizontal lines. This is what occurs in Fig. 4 after about 9000 samples. It should be noted that the network here converges less rapidly than in Fig. 2. This results from two phenomena. On the one

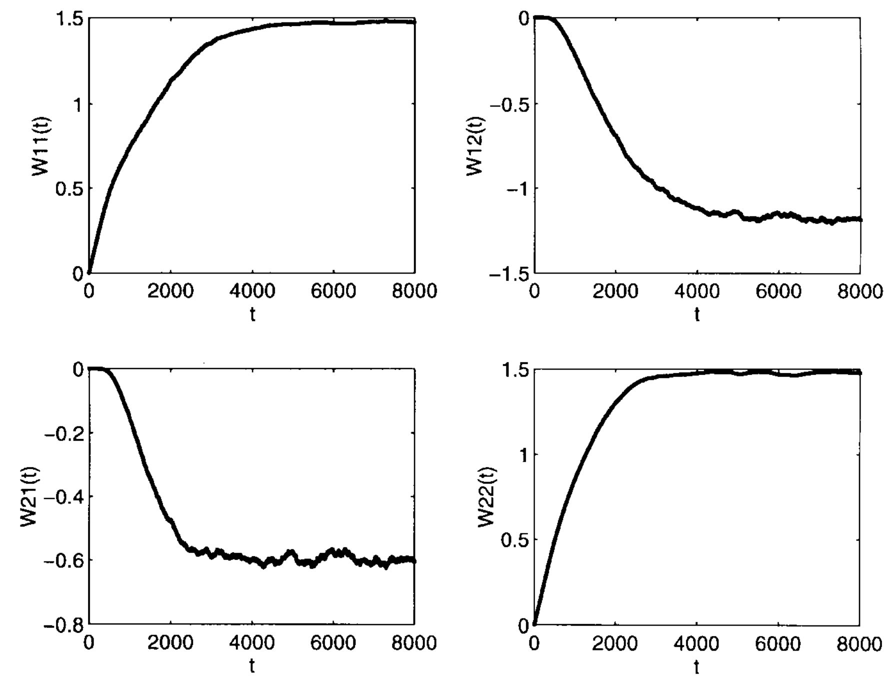


Fig. 3. Network weights for sources corresponding to $p = 0.5$.

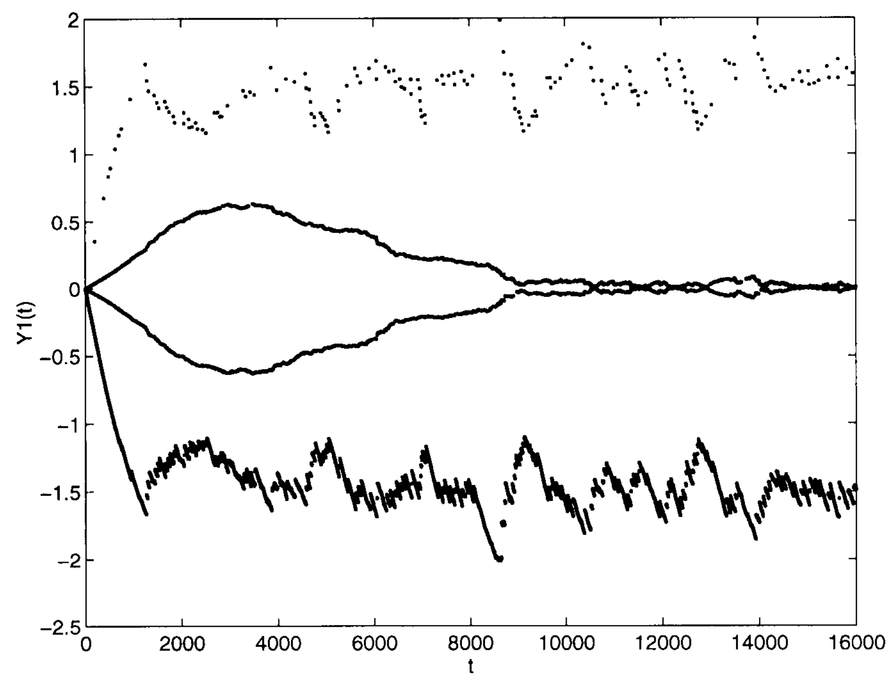


Fig. 4. Network output signal $Y_1(t)$ for sources corresponding to $p = 0.1$ (discrete set of points: see text).

hand, we observed that in the situations where the network eventually converges to a nonseparating point, the weight values always "wander" during a significant period before evolving toward that nonseparating point. On the contrary, convergence toward a separating point occurs directly and, therefore, more rapidly, like in Fig. 2. On the other hand, this low convergence speed may result from the highly asymmetric nature of the sources considered here (i.e., $p \ll 0.5$), which means that a large number of samples must be processed before the network has received the number of samples equal to 1, which is required to achieve convergence. This asymmetric, i.e., "irregular," nature of the sources is also expected to be responsible for the large magnitude of the fluctuations of the weights and, therefore, of $y_1(t)$, as compared with those that appear in Fig. 2. These interpretations are confirmed by the simulations reported in Appendix D.

*3) Conclusions Derived from Simulations:* The simulations described above show that the sources that are separated by the version of the Cichocki network operating with the separating functions (7) are the globally sub-Gaussian signals. This fully confirms the theoretical analysis presented in the previous sections.

## VI. General Conclusions and Prospects

The Cichocki network has been claimed to have significant advantages over the previously published Hérault–Jutten network, but up to now, its exact conditions of operation had not been described. The theoretical analysis and computer simulations reported in this paper prove that the stationary independent sources that can be separated by the standard version of this network are the globally sub-Gaussian signals. As the corresponding Hérault-Jutten network applies to the same sources, this shows that the new features provided by the Cichocki network are not obtained at the expense of a degradation of the field of application. From this point of view, this paper confirms the attractiveness of the Cichocki network on the basis of objective criteria in the considered conditions. To our knowledge, this is the first reported analysis for a four-adaptive-weight source separation neural network, whereas the previous papers concerned networks in which only two weights are adapted (i.e., the diagonal weights of their weight matrix are fixed to 1). It should be remembered, however, that this analysis only applies to a small adaptation gain, as explained in Section IV-A1.

As this paper shows that the standard version of this network cannot separate globally super-Gaussian source signals, we are led to wonder whether a modified version can be defined for such sources. We propose and analyze such a network, based on another set of separating functions, in Appendix D. As a result of our overall investigation, a method for processing each one of the two classes of signals (i.e., sub- and super-Gaussian) is thus available.

This investigation may be extended by analyzing the stability conditions associated with the discrete-time stochastic version of the network adaptation rules. This analysis could be based on the approach developed by Moreau and Macchi for their networks [13], [16]. The behavior of Cichocki networks should also be studied for nonstationary source signals and for other "configurations" than the one defined in Section I. It is likely to be different from that of the Hérault–Jutten networks due to the self-normalization scheme of these new networks.

## Appendix A
## Determination of the Equilibrium Points with Only Nonzero Elements $p_{ij}$

As stated in Section III, the equilibrium points of the considered adaptation rules are all the quadruplets of weight values $(w_{11}, w_{12}, w_{21}, w_{22})$, which meet the set of four equations: $EQ_1 = \{(19), (20), (21), (22)\}$. In this appendix, we determine the solutions of these equations meeting the following requirement:

$$p_{ij} \neq 0, \quad i, j \in \{1, 2\}. \tag{79}$$

Due to (79), the subset $\{(19), (20)\}$ of $EQ_1$ is equivalent to $\{(19), (80)\}$, where (80) is defined as the following linear combination of (19) and (20): $(80) = (19)\ p_{21}^2 p_{22}^2 - (20)\ p_{11}^2 p_{12}^2$. Equation (80) is chosen to get rid of the term $\langle s_1^2\rangle\ \langle s_2^2\rangle$, which occurs in (19) and (20), and reads explicitly

$$\begin{aligned}\left(p_{11}^2p_{22}^2 - p_{12}^2p_{21}^2\right)\left(p_{11}^2p_{21}^2\langle s_1^4\rangle - p_{12}^2p_{22}^2\langle s_2^4\rangle\right)\\ = p_{21}^2p_{22}^2 - p_{11}^2p_{12}^2.\end{aligned} \tag{80}$$

Similarly, due to (79), the subset $\{(21), (22)\}$ of $EQ_1$ is equivalent to $\{(21), (81)\}$, where (81) is defined as the following linear combination of (21) and (22): $(81) = (21)p_{21}p_{22} - (22)p_{11}p_{21}$. Equation (81) is also chosen to get rid of the term $\langle s_1^2\rangle\langle s_2^2\rangle$, which occurs in (21) and (22), and reads explicitly

$$(p_{11}p_{22} - p_{12}p_{21})\left(p_{11}^2p_{21}^2\langle s_1^4\rangle - p_{12}^2p_{22}^2\langle s_2^4\rangle\right) = 0. \tag{81}$$

On assumption (79), the set of equations $EQ_1$ to be solved is thus replaced by the equivalent set $EQ_2 = \{(19), (80), (21), (81)\}$. The latter set of equations can then be simplified by using the following theorems.

*Theorem 4:* At a given point in the weight space, if (79) is met and if $p_{11}p_{22} - p_{12}p_{21} = 0$, this point is not an equilibrium point of the adaptation algorithm considered in this paper.

*Proof:* If (79) is met and if $p_{11}p_{22} - p_{12}p_{21} = 0$, then (16) and (17) yield

$$y_1(t) = \frac{p_{11}}{p_{21}} y_2(t). \tag{82}$$

Hence, we derive

$$\langle [y_1(t)]^3 y_2(t)\rangle = \left(\frac{p_{11}}{p_{21}}\right)^3 \langle [y_2(t)]^4\rangle. \tag{83}$$

If the considered point were an equilibrium point, (11) and (12) would be met and would especially yield

$$\langle [y_2(t)]^4\rangle = 1 \tag{84}$$

$$\langle [y_1(t)]^3 y_2(t)\rangle = 0. \tag{85}$$

As (84) and (85) are not compatible with (83) and (79), we conclude that in the considered conditions, the point is not an equilibrium point. □

*Corollary 1:* If a point in the weight space is an equilibrium point of the adaptation algorithm considered in this paper and if (79) is met, then $p_{11}p_{22} - p_{12}p_{21} \neq 0$.

Hence, we derive that on assumption (79), the set of equations $EQ_2$ is equivalent to the set $EQ_3 = \{(19), (80), (21), (86)\}$, where (86) is defined as[11]

$$p_{11}^2p_{21}^2\langle s_1^4\rangle = p_{12}^2p_{22}^2\langle s_2^4\rangle. \tag{86}$$

Moreover, it is easily shown that assuming (79), the subset $\{(80), (86)\}$ of $EQ_3$ is equivalent to $\{(87), (86)\}$, where (87) is defined as

$$p_{21}^2p_{22}^2 = p_{11}^2p_{12}^2. \tag{87}$$

The set of equations $EQ_3$ is thus replaced by the equivalent set $EQ_4 = \{(19), (87), (21), (86)\}$, which consists of

[11] The equivalence between $EQ_2$ and $EQ_3$, assuming (79), may be shown as follows: If $EQ_2$ is met, then the considered point is an equilibrium point; therefore, Corollary 1 yields $p_{11}p_{22} - p_{12}p_{21} \neq 0$. This term can then be simplified in (81), thus yielding (86) and, therefore, $EQ_3$. Conversely, $EQ_3$ always implies $EQ_2$.

a subset $\{(19),(21)\}$ of complicated equations and a subset $\{(87),(86)\}$ of simpler equations. As a first step, we solve the latter subset of equations, i.e., we use it to express $p_{21}$ and $p_{22}$ with respect to $p_{11}, p_{12}$ and the statistics of the source signals. We thus derive an equivalent subset obtained by multipliying (87) and (86) either term by term or in a crossed way

$$p_{21}^4\langle s_1^4\rangle = p_{12}^4\langle s_2^4\rangle \tag{88}$$
$$p_{11}^4\langle s_1^4\rangle = p_{22}^4\langle s_2^4\rangle. \tag{89}$$

The solution of these equations (for real-valued terms $p_{ij}$) reads

$$p_{21} = \epsilon_1 p_{12} R, \quad \text{with } \epsilon_1 \in \{-1,1\} \tag{90}$$
$$p_{22} = \frac{\epsilon_1' p_{11}}{R}, \quad \text{with } \epsilon_1' \in \{-1,1\} \tag{91}$$

where the ratio $R$ is defined as

$$R = \left(\frac{\langle s_2^4\rangle}{\langle s_1^4\rangle}\right)^{1/4}. \tag{92}$$

The set of equations $EQ_4$ is thus replaced by the equivalent set $EQ_5 = \{(19),(21),(90),(91)\}$. Equations (90) and (91) then allow us to replace $p_{21}$ and $p_{22}$ in (21), thus yielding (still with $\epsilon_1 \in \{-1,1\}$ and $\epsilon_1' \in \{-1,1\}$)

$$\begin{aligned}&\left(R^2\langle s_1^4\rangle + 3\epsilon_1\epsilon_1'\langle s_1^2\rangle\langle s_2^2\rangle\right)p_{11}^2\\ &\quad + \left(\epsilon_1\epsilon_1'\langle s_2^4\rangle + 3R^2\langle s_1^2\rangle\langle s_2^2\rangle\right)p_{12}^2 = 0.\end{aligned} \tag{93}$$

Moreover, due to the signs of its terms, the latter equation has a solution only when

$$\epsilon_1' = -\epsilon_1 \tag{94}$$

for which it becomes

$$\begin{aligned}&\left(R^2\langle s_1^4\rangle - 3\langle s_1^2\rangle\langle s_2^2\rangle\right)p_{11}^2\\ &\quad + \left(-\langle s_2^4\rangle + 3R^2\langle s_1^2\rangle\langle s_2^2\rangle\right)p_{12}^2 = 0.\end{aligned} \tag{95}$$

As stated above, (21) can thus be replaced by (95) in $EQ_5$. In addition, (94) allows us to replace (91) by

$$p_{22} = \frac{-\epsilon_1 p_{11}}{R}. \tag{96}$$

The set of equations $EQ_5$ is thus replaced by the equivalent set $EQ_6 = \{(19),(95),(90),(96)\}$. Moreover, (95) may be rewritten as

$$Q\left(p_{11}^2 - R^2 p_{12}^2\right) = 0 \tag{97}$$

with

$$Q = \sqrt{\langle s_1^4\rangle\langle s_2^4\rangle} - 3\langle s_1^2\rangle\langle s_2^2\rangle \tag{98}$$

so that the set $EQ_6$ of equations is equivalent to $EQ_7 = \{(19),(97),(90),(96)\}$. This shows that the considered algorithm yields a specific case, i.e., $Q = 0$, which corresponds to globally Gaussian sources (defined in Section IV). In this case, (97) is always fulfilled and disappears. $EQ_7$ then becomes a set of only three equations with four unknowns (all terms $p_{ij}$), which has an infinite number of solutions. In other words, in this case, the algorithm has an infinite number of equilibrium points that do not achieve source separation. In the remainder of this appendix, we only consider the case when the sources are not globally Gaussian, i.e., $Q \neq 0$. Equation (97) is then equivalent to

$$p_{11}^2 - R^2 p_{12}^2 = 0 \tag{99}$$

so that the set $EQ_7$ of equations to be solved is equivalent to $EQ_8 = \{(19),(99),(90),(96)\}$. Solving (99) yields the expression of $p_{12}$ versus $p_{11}$, i.e.,

$$p_{12} = \frac{\epsilon_2 p_{11}}{R}, \quad \epsilon_2 \in \{-1,1\}. \tag{100}$$

Inserting (100) in (19) yields

$$p_{11} = \frac{\epsilon_3}{\left[2\langle s_1^4\rangle\left(1 + 3\frac{\langle s_1^2\rangle\langle s_2^2\rangle}{\sqrt{\langle s_1^4\rangle\langle s_2^4\rangle}}\right)\right]^{1/4}}, \quad \epsilon_3 \in \{-1,1\}. \tag{101}$$

The set $EQ_8$ of equations is thus equivalent to $EQ_9 = \{(101),(100),(90),(96)\}$. Equation (100) may eventually be inserted in (90) in order to express all terms $p_{ij}$ with respect to $p_{11}$, which is defined explicitly in (101). $EQ_9$ is thus replaced by a set of four equations, which provide the final solution of the considered problem and that read explicitly

$$p_{11} = \frac{\epsilon_3}{\left[2\langle s_1^4\rangle\left(1 + 3\frac{\langle s_1^2\rangle\langle s_2^2\rangle}{\sqrt{\langle s_1^4\rangle\langle s_2^4\rangle}}\right)\right]^{1/4}}, \quad \epsilon_3 \in \{-1,1\} \tag{102}$$
$$p_{12} = \frac{\epsilon_2 p_{11}}{R}, \quad \epsilon_2 \in \{-1,1\} \tag{103}$$
$$p_{21} = \epsilon_1\epsilon_2 p_{11}, \quad \text{with } \epsilon_1 \in \{-1,1\} \tag{104}$$
$$p_{22} = \frac{-\epsilon_1 p_{11}}{R} \tag{105}$$

where $R$ is defined in (92).

## Appendix B
## Proof of Theorem Defining a Specific Type of Separating Equilibrium Point

In this appendix, we restrict ourselves to the case when the separating functions (7) are used,[12] and we provide the proof of Theorem 2 of Section IV-E.

For any of the equilibrium points $F_{\epsilon_1,\epsilon_2}$, (18) and (23)–(26) yield

$$w_{11} = \frac{a_{22}}{C}p_{11} \tag{106}$$
$$w_{12} = \frac{-a_{12}}{C}p_{11} \tag{107}$$
$$w_{21} = \frac{-a_{21}}{C}p_{22} \tag{108}$$
$$w_{22} = \frac{a_{11}}{C}p_{22} \tag{109}$$

[12] On the opposite of the situation of Theorem 3, there is no need to require the sources to be globally sub-Gaussian here.

with

$$C = a_{11}a_{22} - a_{12}a_{21}. \tag{110}$$

It should be noted that the hypothesis $a_{11}a_{22} - a_{12}a_{21} \neq 0$ mentioned in Section I is equivalent to $C \neq 0$. Equations (106)–(109) show that this condition is required for the equilibrium points $F_{\epsilon_1,\epsilon_2}$ to exist (with finite weight values).

For given nonzero mixture coefficients $a_{ij}$, (23) and (106) show that $w_{11}$ is positive or negative, depending on $\epsilon_1$. Similarly, (26) and (109) show that $w_{22}$ is positive or negative, depending on $\epsilon_2$. Therefore, exactly one of the four possible couples $(\epsilon_1, \epsilon_2)$, i.e., exactly one of the four points $F_{\epsilon_1,\epsilon_2}$, is such that $w_{11}$ and $w_{22}$ are positive.

Similarly, for any of the equilibrium points $G_{\epsilon_1,\epsilon_2}$, (18) and (29)–(32) yield

$$w_{11} = \frac{-a_{21}}{C}p_{12} \tag{111}$$
$$w_{12} = \frac{a_{11}}{C}p_{12} \tag{112}$$
$$w_{21} = \frac{a_{22}}{C}p_{21} \tag{113}$$
$$w_{22} = \frac{-a_{12}}{C}p_{21} \tag{114}$$

so that exactly one of the four points $G_{\epsilon_1,\epsilon_2}$ is such that $w_{11}$ and $w_{22}$ are positive.

Now, consider the corresponding values of $D$, which are defined in (45). For any of the equilibrium points $F_{\epsilon_1,\epsilon_2}$, (106)–(109) yield

$$D = \frac{C}{a_{11}a_{22}}w_{11}w_{22}. \tag{115}$$

Similarly, for any of the equilibrium points $G_{\epsilon_1,\epsilon_2}$, (111)–(114) yield

$$D = \frac{-C}{a_{12}a_{21}}w_{11}w_{22}. \tag{116}$$

It should first be noted that all these values of $D$ are nonzero since $C$, $w_{11}$, and $w_{22}$ are themselves nonzero, as explained above. Now, consider the two points among $F_{\epsilon_1,\epsilon_2}$ and $G_{\epsilon_1,\epsilon_2}$ that are such that $w_{11}$ and $w_{22}$ are positive and assume that the two corresponding values of $D$ were negative. Equations (115) and (116) would then yield

$$\frac{C}{a_{11}a_{22}} < 0 \quad \text{and} \quad \frac{-C}{a_{12}a_{21}} < 0 \tag{117}$$

or equivalently

$$\frac{a_{11}a_{22}}{C} < 0 \quad \text{and} \quad \frac{-a_{12}a_{21}}{C} < 0 \tag{118}$$

and therefore, by adding the latter two expressions

$$\frac{a_{11}a_{22} - a_{12}a_{21}}{C} < 0 \tag{119}$$

i.e., $1 < 0$. Because this is not true, at least one of the two points among $F_{\epsilon_1,\epsilon_2}$ and $G_{\epsilon_1,\epsilon_2}$ that are such that $w_{11}$ and $w_{22}$ are positive is, in addition, such that $D$ is positive. This proves Theorem 2.

## Appendix C
## Proof of Theorem Defining a Stable Separating Equilibrium Point

In this appendix, we restrict ourselves to the case when the separating functions (7) are used and the sources are globally sub-Gaussian, and we provide the proof of Theorem 3 of Section IV-E. To this end, we first consider a specific case for which we derive a restricted version of Theorem 3, and we then investigate the general case of interest.

### *A. Theorem and Proof for a Specific Case*

The type of separating equilibrium point(s) to be eventually considered is the one defined in Theorem 2, i.e., the point(s) such that $w_{11} > 0$, $w_{22} > 0$, and $D > 0$. Before considering all such points in subsection B of this Appendix, we here focus on a subset of these points, consisting of the points that are such that $w_{11} > 0$, $w_{22} > 0$, and $w_{12}w_{21} = 0$ (these conditions entail $D > 0$). For such points, the following theorem holds.

*Theorem 5:* For the separating functions (7) and globally sub-Gaussian sources, if a separating equilibrium point is such that $w_{11} > 0$, $w_{22} > 0$, and $w_{12}w_{21} = 0$, it is stable.

*Proof:* When setting the condition $w_{12}\, w_{21} = 0$ (without any additional conditions on $w_{11}$ and $w_{22}$), the expressions of the four roots $\lambda_i$ of the polynomial $P(\lambda)$ defined in (49)–(54) become much simpler, i.e.,

$$\lambda_1 = 4w_{11} \tag{120}$$
$$\lambda_2 = 4w_{22} \tag{121}$$
$$\lambda_3 = \frac{(w_{11}+w_{22})3\langle y_1^2 y_2^2\rangle + \Delta^{1/2}}{2} \tag{122}$$
$$\lambda_4 = \frac{(w_{11}+w_{22})3\langle y_1^2 y_2^2\rangle - \Delta^{1/2}}{2} \tag{123}$$

with

$$\Delta = 9\langle y_1^2 y_2^2\rangle^2 (w_{11}-w_{22})^2 + 4w_{11}w_{22}. \tag{124}$$

Moreover, it is here assumed that $w_{11} > 0$, $w_{22} > 0$, and the sources are globally sub-Gaussian, i.e., (57) is met for any separating equilibrium point. In this case, we derive easily from (120)–(123) that all four roots $\lambda_i$ of the polynomial $P(\lambda)$ are real and positive. Moreover, as $D = w_{11}w_{22} > 0$, (44) yields $r < 0$. All four roots of the polynomial $Q(\Lambda)$ are therefore real and negative, due to (55). Condition C2) of Section IV-A1 then entitles us to conclude that any such separating equilibrium point is stable.

### *B. Proof for the General Case*

In this subsection, we provide a proof of Theorem 3, based on an extension of Theorem 5 above. To this end, we investigate the stability of an arbitrary separating equilibrium point, which is assumed to be such that $w_{11}^T > 0$, $w_{22}^T > 0$, and $D^T > 0$, where a superscript "$T$" is used in the notations related to the weights $w_{ij}^T$ and to the parameter $D^T$ defined in (45), in order to indicate that these values correspond to the point where stability is to be tested. The approach used below consists of linking the stability of this test point to the stability of a reference point, which is a separating equilibrium point corresponding to other values of the mixture coefficients $a_{ij}$. The latter coefficients are chosen so that the reference point has the same values as the test point for the parameters $w_{11}$ and $w_{22}$ but is such that $w_{12}w_{21} = 0$. Especially, we hereafter consider the reference point defined as follows, where a superscript "$R$" is used in the notations related to the weights $w_{ij}^R$ and to the parameter $D^R$ of this reference point

$$w_{11}^R = w_{11}^T \tag{125}$$
$$w_{12}^R = w_{12}^T \tag{126}$$
$$w_{21}^R = 0 \tag{127}$$
$$w_{22}^R = w_{22}^T. \tag{128}$$

In order to link these two points, we consider a continuous trajectory from the test point to the reference point in the weight space, which corresponds to varying the coefficients $a_{ij}$ so that $w_{21}$ at the considered separating equilibrium point is varied from $w_{21}^T$ to $w_{21}^R = 0$. Each point of this trajectory is, therefore, an intermediate point (leading to a superscript "$I$" in the notations used for its weights $w_{ij}^I$ and for the parameter $D^I$) between the test point and the reference point, which is defined as

$$w_{11}^I = w_{11}^T \tag{129}$$
$$w_{12}^I = w_{12}^T \tag{130}$$
$$w_{21}^T \quad \text{varied from } w_{21}^T \text{ to } w_{21}^R = 0 \tag{131}$$
$$w_{22}^I = w_{22}^T. \tag{132}$$

As it is assumed that $w_{11}^T > 0$ and $w_{22}^T > 0$, (129) and (132) yield $w_{11}^I > 0$, and $w_{22}^I > 0$. Moreover, $D^I = w_{11}^T w_{22}^T - w_{12}^T w_{21}^I$ is a linear function of $w_{21}^I$ and is therefore comprised of $D^T > 0$ and $D^R = w_{11}^T \ w_{22}^T > 0$. Therefore, $D^I > 0$. The conditions $w_{11}^I > 0$, $w_{22}^I > 0$, and $D^I > 0$, combined with the fact that the sources are assumed to be sub-Gaussian [i.e., (57) is met], imply that at the considered intermediate point, $p_1$ and $p_3$, respectively, defined in (51) and (53), are both negative. Therefore, $p_1/p_3 > 0$. It can be shown that the latter condition, combined with the fact that $w_{11}^I$, $w_{22}^I > 0$, $D^I > 0$, and (57) imply that none of the roots of the polynomial $P(\lambda)$ corresponding to (50)–(54) has a zero real part.[13] As this applies to any intermediate point, it means that when the considered intermediate point is varied continuously from the test point to the reference point, the real parts of the roots of $P(\lambda)$ never become equal to zero. As these roots vary continuously, this implies that their sign remains constant and, therefore, equal to the sign that they have for the reference point, i.e., positive (see Theorem 5 and its proof). Moreover, since $D^I > 0$, (44) yields $r < 0$ for any intermediate point. Therefore, all four roots of the polynomial $Q(\Lambda)$ have a negative real part, due to (55). Condition C2) of Section IV-A1 then allows us to conclude that any such considered intermediate point is stable. The test point is especially stable, which yields Theorem 3.

[13] This result is obtained by determining the purely imaginary solutions of $P(\lambda) = 0$, i.e., the solutions with $\lambda = ix$, where $x$ is a real unknown. This yields a set of first- and second-order equations, which has no solution for the considered polynomial, due to the above-mentioned conditions. For the sake of brevity, the simple but somewhat long resolution of this set of equations is omitted here.

## APPENDIX D
## DEFINITION OF A NETWORK SUITED TO SUPER-GAUSSIAN SOURCES

In this appendix, we are concerned with the determination of a set of separating functions $f$ and $g$ resulting in a Cichocki network that is able to separate the sources that cannot be processed by the standard network considered above. In other words, we look for a version of this network suited to super-Gaussian sources. To this end, we take advantage of the similarity of the properties shown above between the standard versions of the Hérault–Jutten and Cichocki networks, i.e., the versions that operate with the separating functions defined in (7) and are called the $(3, 1)$ networks hereafter. Based on this similarity, we now consider a modified version of the Cichocki network, which operates with

$$f(x) = x \quad \text{and} \quad g(x) = x^3 \tag{133}$$

and is called the $(1, 3)$ Cichocki network hereafter. We expect this modified network to apply to the same type of sources as the corresponding $(1, 3)$ Hérault–Jutten network, i.e., to super-Gaussian sources.[14]. The remainder of this appendix aims at checking that this conjecture is true. The method used to this end is the same as in the previous sections. Therefore, only its aspects specific to the $(1, 3)$ network are detailed hereafter.

### A. Equilibrium Points

Whatever the selected couple of separating functions $(f, g)$, the equilibrium conditions for the corresponding version of the Cichocki network are (9) and (10). Applying them to the $(1, 3)$ network based on the separating functions (133) and rearranging them, we again get (11) and (12). Therefore, the $(1, 3)$ network considered in this appendix yields exactly the same equilibrium points as the $(3, 1)$ network analyzed above. We then conclude, in the same way as in Section III-D, that the stability of these equilibrium points with respect to the modified adaptation rule considered in this appendix must be studied.

### B. Stability of the Equilibrium Points

*1) Tangent Mean Algorithm and First Stability Conditions:* The local stability of the equilibrium points of the considered network is analyzed by applying the approach defined in Section IV-A1 to the separating functions (133). This again yields (41)–(45) and the associated stability conditions C1)

[14] This result concerning the $(1, 3)$ Hérault–Jutten network may be derived from [10] and is explicitly provided in [4] and [5].

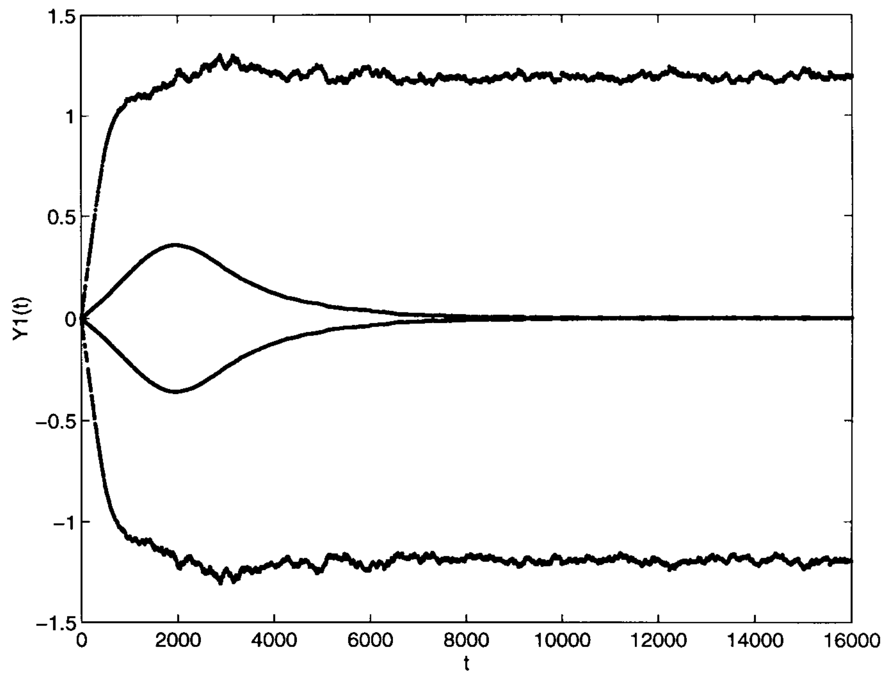


Fig. 5. Output signal $Y_1(t)$ of the $(1, 3)$ network for sources corresponding to $p = 0.5$ (discrete set of points: see text).

and C2), but with

$$M = \begin{bmatrix} 4w_{22} & -4w_{21} & 0 & 0 \\ -w_{12} & w_{11} & 3w_{22}\langle y_1^2 y_2^2 \rangle & -3w_{21}\langle y_1^2 y_2^2 \rangle \\ -3w_{12}\langle y_1^2 y_2^2 \rangle & 3w_{11}\langle y_1^2 y_2^2 \rangle & w_{22} & -w_{21} \\ 0 & 0 & -4w_{12} & 4w_{11} \end{bmatrix}. \quad (134)$$

From this, we again derive (47)–(49), but with

$$p_4 = 1 \quad (135)$$

$$p_3 = -5(w_{11} + w_{22}) \quad (136)$$

$$p_2 = -\langle y_1^2 y_2^2 \rangle^2 9 w_{11} w_{22} + (w_{11} + 4w_{22})(w_{22} + 4w_{11}) + 8D \quad (137)$$

$$p_1 = 4(w_{11} + w_{22})D\left[9\langle y_1^2 y_2^2 \rangle^2 - 5\right] \quad (138)$$

$$p_0 = 16D^2\left[1 - 9\langle y_1^2 y_2^2 \rangle^2\right]. \quad (139)$$

The corresponding new polynomial $P(\lambda)$ is therefore used as an intermediate variable in the following analysis, which is again based on the original approach introduced in Section IV-B.

*2) Necessary Stability Condition:* As shown in Section IV-C, Theorem 1 applies to any fourth-order polynomial $P(\lambda)$ with $p_4 = 1$ and, therefore, to any associated set of separating functions. When applied to the separating functions (133) considered here and to the corresponding polynomial (49), (135)–(139), it yields the necessary stability condition

$$\langle y_1^2 y_2^2 \rangle < \frac{1}{3}. \quad (140)$$

Using the expressions of $\langle y_1^2 y_2^2 \rangle$ with respect to the source signals that were derived in Section IV for each equilibrium point, (140) may be rewritten as follows:

- For any of the separating equilibrium points $F_{\epsilon_1,\epsilon_2}$ and $G_{\epsilon_1,\epsilon_2}$, (140) is equivalent to (61).
- For any of the nonseparating equilibrium points $H_{\epsilon_1,\epsilon_2,\epsilon_3}$, (140) is equivalent to (59).

*3) Interpretation of the Necessary Stability Condition:* Using the same approach as in Section IV-D, the above conditions allow us to conclude the following:

- The $(1, 3)$ network cannot separate globally sub-Gaussian and Gaussian sources. This completes our stability analysis for such sources. The remainder of our analysis therefore only concerns the other type of sources.
- For globally super-Gaussian sources, "situation1" and "situation2," defined in Section IV-D, are again *a priori* possible at this stage of the investigation. Additional properties are therefore considered hereafter to determine which of these situations actually occurs.

*4) Definition of a Stable Separating Equilibrium Point:* As explained above, we here restrict ourselves to the case when the separating functions (133) are used and the sources are globally super-Gaussian. The last step of the stability analysis is then based on the following theorems.

*Theorem 6:* For the separating functions (133), if all mixture coefficients $a_{ij}$ are nonzero, at least one of the separating equilibrium points is such that $w_{11} > 0$, $w_{22} > 0$, and $D > 0$, where $D$ is defined in (45).

*Theorem 7:* For the separating functions (133) and for globally super-Gaussian sources, if a separating equilibrium point is such that $w_{11} > 0$, $w_{22} > 0$, and $D > 0$, it is stable.

A theorem similar to Theorem 6 was only proved for the $(3, 1)$ network operated with the separating functions (7) in Section IV. However, it also directly applies to the $(1, 3)$ network adapted with the functions (133), as the proof of this theorem only depends on the expressions of the equilibrium

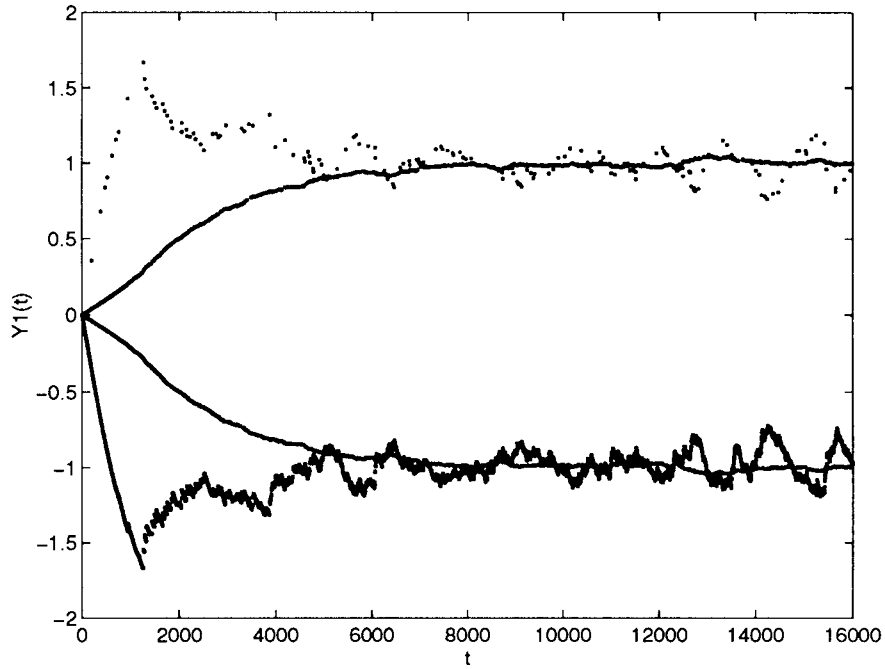


Fig. 6. Output signal $Y_1(t)$ of the $(1, 3)$ network for sources corresponding to $p = 0.1$ (discrete set of points: see text).

points, which are the same for both networks. As for Theorem 7, a new demonstration is required for the network considered in this appendix. This demonstration is not detailed here because it is based on the same principles as that of the corresponding theorem of Section IV.[15]

From these theorems, we derive, in the same way as in Section IV-E, that the $(1, 3)$ network is able to separate all the globally super-Gaussian sources (and none of the other types of sources as was proved above), which completes the stability analysis.

### C. Simulation Results for the $(1, 3)$ Network

The same type of simulations as in Section V was also used to validate all the results derived above for the $(1, 3)$ network. These simulations were again performed for various values of the probability $p$ defining the sources. The results thus obtained are symmetric to those provided in Section V and are therefore described more briefly hereafter.

With sub-Gaussian sources, i.e., when $p$ is chosen so that (65) is met, the network converges to a nonseparating equilibrium point. This is illustrated in Fig. 5 for the case $p = 0.5$. The convergence phase lasts about 7000 samples. This value is comprised between those observed in the previous two cases, i.e., 4000 and 9000 samples. This results from the fact that only one of the two speed-reduction phenomena defined in Section V-B2 occurs here. The network again "wanders" before eventually reaching the nonseparating point, but it does not suffer from any source asymmetry. After convergence, the output signal $Y_1(t)$ again has three possible values (corresponding to the three horizontal lines in the figure), showing that the network converges to a nonseparating point. Moreover, the magnitude of the weight fluctuations is low, as the sources are symmetric.

With super-Gaussian sources, i.e. when $p$ is chosen so that (68) is met, the network converges to a separating equilibrium point. This is illustrated in Fig. 6 for the case where $p = 0.1$. The asymmetric nature of the sources is again responsible for an intermediate convergence speed (7000 samples) and for the fluctuations of the weights. After convergence, the output signal $Y_1(t)$ only has two possible values (corresponding to the two horizontal lines in the figure).

Therefore, the sources that are separated by the $(1, 3)$ network are the globally super-Gaussian signals. This fully confirms the theoretical analysis presented in the previous sections of this appendix.

[15] When $w_{12}\ w_{21} = 0$, $P(\lambda)$ has the same first two roots as in Section IV, whereas its last two roots become

$$\lambda_3 = \frac{(w_{11} + w_{22}) + \Delta^{1/2}}{2} \quad \text{and} \quad \lambda_4 = \frac{(w_{11} + w_{22}) - \Delta^{1/2}}{2} \tag{141}$$

with

$$\Delta = (w_{11} - w_{22})^2 + 36 w_{11} w_{22} \langle y_1^2 y_2^2 \rangle^2. \tag{142}$$

## Acknowledgment

The author would like to thank A. Deville and N. Charkani for their useful comments about the polynomials considered in the stability analysis presented in this paper. He also wishes to thank L. Andry, G. Boyer, and A. Cichocki for helpful discussions about simulation results.

## References

[1] C. Jutten and J. Hérault, "Blind separation of sources, part I: An adaptive algorithm based on neuromimetic architecture," *Signal Process.*, vol. 24, no. 1, pp. 1–10, July 1991.

[2] Y. Bar-Ness, J. W. Carlin, and M. L. Steinberger, "Bootstrapping adaptive cross pol cancelers for satellite communications," in *Proc. IEEE Int. Conf. Commun. ICC*, Miami, FL, June 1982, pp. 4F.5.1–4F.5.5.

[3] E. Chaumette and D. Muller, "Séparation de sources aux ordres supérieurs: Application au garbling en radar secondaire," in *Proc. Ecole Techniques Avancées Signal Image Parole*, Grenoble, France, Sept. 2–6, 1996, pp. 245–253.

[4] Y. Deville, "Application of the Hérault-Jutten source separation neural network to multi-tag radio-frequency identification systems" (invited paper), in *Proc. Ecole Techniques Avancées Signal Image Parole*, Grenoble, France, Sept. 2–6, 1996, pp. 265–272.

[5] Y. Deville and L. Andry, "Application of blind source separation techniques to multi-tag contactless identification systems," *IEICE Trans. Fundamentals Electron., Commun. Comput. Sci.*, vol. E79-A, no. 10, pp. 1694–1699, Oct. 1996.

[6] G. D'Urso, "Les techniques de séparation de sources appliquées à la surveillance des instalations EDF," in *Proc. Ecole Techniques Avancées Signal Image Parole*, Grenoble, France, Sept. 2–6, 1996, pp. 255–263.

[7] J. Hérault and B. Ans, "Circuits neuronaux à synapses modifiables: Décodage de messages composites par apprentissage nonsupervisé," *C. R. Acad. Sci. Paris*, série III, no. 13, 1984.

[8] C. Jutten and J. Hérault, "Une solution neuromimétique au problème de séparation de sources," *Traitement Signal*, vol. 5, no. 6, pp. 389–403, 1988.

[9] K. J. Pope and R. E. Bogner, "Blind signal separation I. Linear, instantaneous combinations," *Digital Signal Process.*, vol. 6, pp. 5–16, 1996.

[10] E. Sorouchyari, "Blind separation of sources, part III: Stability analysis," *Signal Process.*, vol. 24, no. 1, pp. 21–29, July 1991.

[11] J.-C. Fort, "Stabilité de l'algorithme de séparation de sources de Jutten et Hérault," *Traitement Signal*, vol. 8, no. 1, pp. 35–42, Jan. 1991.

[12] E. Moreau and O. Macchi, "Self-adaptive source separation, part II: Comparison of the direct, feedback and mixed neural network," submitted for publication.

[13] E. Moreau, "Apprentissage et adaptativité. Séparation auto-adaptative de sources indépendantes," Ph.D. dissertation, Univ. Paris XI, Orsay, France, Feb. 1, 1995.

[14] P. Comon, C. Jutten, and J. Hérault, "Blind separation of sources, part II: Problems statement," *Signal Process.*, vol. 24, no. 1, pp. 11–20, 1991.

[15] Y. Deville, "A unified stability analysis of the Hérault-Jutten source separation neural network," *Signal Process.*, vol. 51, no. 3, pp. 229–233, June 1996.

[16] O. Macchi and E. Moreau, "Self-adaptive source separation, part I: Convergence analysis of a direct linear network controled by the Hérault-Jutten algorithm," *IEEE Trans. Signal Processing*, vol. 45, pp. 918–926, Apr. 1997.

[17] A. Cichocki, W. Kasprzak, and S. I. Amari, "Multi-layer neural networks with a local adaptive learning rule for blind separation of source signals," in *Proc. 1995 Int. Symp. Nonlinear Theory Appl. NOLTA*, Las Vegas, NV, Dec. 10–14, 1995, pp. 61–65.

[18] A. Benveniste, M. Metivier, and P. Priouret, "Adaptive algorithms and stochastic approximations," *Appl. Math.*, vol. 22, 1990.

[19] R. Gooch and J. Lundell, "The CM array: An adaptive beamformer for constant modulus signals," in *Proc. Int. Conf. Acoust., Speech Signal Process.*, Tokyo, Japan, Apr. 7–11, 1986, pp. 2523–2526.

[20] E. Durand, *Solutions Numériques des Équations Algébriques, Tome I.* Paris, France: Masson, 1960.

[21] A. Korganoff, *Méthodes de Calcul Numérique, Tome I* Paris, France: Dunod, 1961.

[22] P. I. Richards, *Manual of Mathematical Physics*. London, U.K.: Pergamon, 1959.



**Yannick Deville** (M'90) was born in Lyon, France, on August 10, 1964. He graduated from the Ecole Nationale Supérieure des Télécommunications de Bretagne, Brest, France, in 1986. He received the D.E.A. and Ph.D. degrees, both in microelectronics, from the University of Grenoble, Grenoble, France, in 1986 and 1989, respectively.

From 1986 to 1997, he was with the Laboratoires d'Electronique Philips, Limeil, France, where he was a Senior Research Scientist. His investigations during this period concerned various fields, including GaAs integrated microwave RC active filters, VLSI cache memory replacement algorithms and architectures, neural network algorithms and applications, and nonlinear systems. Since 1997, he has been a Professor at the University of Toulouse, Toulouse, France. His current research interests include signal processing, neural networks, and, especially, blind source separation methods and their acoustic and electromagnetic applications.